(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平7－311833

(43)公開日　平成７年(1995)11月28日

(51)Int.Cl.[6]	識別記号	庁内整理番号	Ｆ Ｉ	技術表示箇所
Ｇ０６Ｔ　1/00				
Ｇ０１Ｂ　11/24		Ｋ		
Ｇ０６Ｔ　9/20				
			Ｇ０６Ｆ　15/ 62　３８０	
		7459－5Ｌ	15/ 70　３３５　Ａ	

審査請求　有　請求項の数４　ＯＬ　（全７頁）　最終頁に続く

(21)出願番号　特願平6－102412

(22)出願日　平成６年(1994)５月17日

(71)出願人　000004237
日本電気株式会社
東京都港区芝五丁目７番１号

(72)発明者　坂本　静生
東京都港区芝五丁目７番１号　日本電気株式会社内

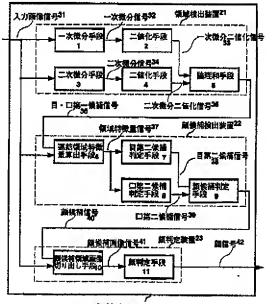
(74)代理人　弁理士　京本　直樹　（外２名）

(54)【発明の名称】　人物の顔の検出装置

(57)【要約】

【目的】　顔の大きさの変化や平行・回転移動、また照明条件の変動に強い、顔検出装置を提供する。

【構成】　人物の顔を含む画像を入力し、顔の構成要素を領域として切り出す領域検出装置２１と、顔構成要素領域の大きさや位置関係より顔の候補を出力する顔候補検出装置２２と、顔候補を詳しく調べて顔であるか否かを判定する顔判定装置２３とから成る。



1

【特許請求の範囲】

【請求項1】人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗そして明に変化する点と、暗から明に変化する点で挟まれる領域を検出する手段と、

該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、

上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、

上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを取り出して顔候補領域として検出する手段とを有することを特徴とする顔検出装置。

【請求項2】人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが暗から明に変化する点と、明から暗そして明に変化する点で挟まれる領域を検出する手段と、

該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、

上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、

上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを取り出して顔候補領域として検出する手段とを有することを特徴とする顔検出装置。

【請求項3】人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗そして明に変化する点と、暗から明に変化する点で挟まれる領域を検出する手段と、

該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、

上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、

上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを顔候補領域として取り出す手段と、

標準顔と上記顔候補領域を比較することにより人物の顔を検出する手段とを有することを特徴とする顔検出装置。

【請求項4】人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗に変化する点と、明から暗そして明に変化する点で挟まれる領域を検出する手段と、

該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、

上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段

2

と、

上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを顔候補領域として取り出す手段と、標準顔と上記顔候補領域を比較することにより人物の顔を検出する手段とを有することを特徴とする顔検出装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は画像中から人物の顔を検出する技術であり、ビルの監視システムやロボットの視覚システムに利用可能である。

【0002】

【従来の技術】人物画像から、人物の顔を検出する従来方式の一例として、文献『自然感の高いビジュアルヒューマンインターフェースの実現のための人物動画像の実時間並列協調的認識（長谷川修、横沢一彦、石塚満、電子情報通信学会論文誌Ｖｏｌ．Ｊ７７－Ｄ－ＩＩ　Ｎｏ．１　Ｊａｎ（１９９４）』に基づいて大まかに説明する。

【0003】該文献で用いられているアルゴリズムは動画像が入力された時、ボトムアップ処理として次に示す特徴を抽出する。

【0004】1．画像中４００点（２０×２０点）の平均輝度

2．画像中２５００点（５０×５０点）のフレーム間差分に基づく動きベクトル

3．画像中２５００点（５０×５０点）の各点における、ＲＧＢ値の比に基づく髪と肌色の領域

4．画像中２５００点（５０×５０点）と設定（背景）画像の間の差分が大きな領域

5．画像中に設定したウインド内のエッジ特徴

ここで特徴抽出処理3、4は閾値の設定が必要となる。固定の閾値を与えると適応できる場面が少なくなるため、処理1の結果によって画像全体の明るさを判別し、これらの閾値を適応的に変化させている。

【0005】処理3では、典型的な日本人が持つ黒い髪と肌色を想定した閾値を設定している。処理3と4の結果は論理積がとられ、設定画像中に存在しない肌色領域を抽出する。更に各領域を動き情報と対応付け、動いている顔を検出する際に用いる。

【0006】また他方でトップダウン処理として顔モデルを以下の条件を用いて構築している。

【0007】1．髪領域と肌領域の面積比が±３０％以内である。

【0008】2．肌色領域上にエッジ情報（目・鼻・口）が設定した範囲の量だけ存在する。

【0009】3．上記エッジ情報は一様な分布をしていない。

【0010】4．エッジ情報は肌色領域上部に比較的集

3

まっている（日）。

本処理はボトムアップ処理の結果顔が存在すると思われる領域について行われ、検出結果はボトムアップ処理中の5を行うためのウインド設定に使用される。

【0011】従来技術の一例としても一つ、『Target image extraction for face recognition using the subspace classification method（S. Akamatsu, T. Sasaki, H. Fukamachi, Y. Suenaga, MVA'92 pp. 465－468（1992））』に基づいて大まかに説明する。

【0012】日本人人物顔が入っているカラー画像が入力された時、該画像をＨＳＶ色空間とＹＩＱ色空間それぞれで観察することにより、次の知見が得られたとしている。

【0013】・ＹＩＱ色空間のＱ軸上で、暑は同りと比較して高い値を持つ。

【0014】・肌領域はＨＳＶ色空間のＨ軸上と、ＹＩＱ色空間のＩ軸上で特徴的なピークを持っている。そこでＲＧＢで入力されたカラー画像をＨＳＶとＹＩＱに色変換して閾値処理を行い肌領域を検出した後、領域の上記特徴により暑領域を検出する。また肌領域内の穴を目候補として、目・口が顔らしい配置を成している組み合わせを顔候補とする。

【0015】得られた顔候補は目と眉を間違えて検出したものも含まれる。図2に検出される顔候補を、目と口で位置の正規化を行い6つに分類して示した。（a）は検出成功した場合に得られる画像であり、（b）と（c）はそれぞれ目の片側を眉であると間違った例である。（d）・（e）・（f）はそれぞれ眼鏡をかけた人の例である。

【0016】本従来技術ではこれら6つの標準画像を用意しておき、（a）または（d）に最も近い候補を検出した顔であるとして出力する。

【0017】

【発明が解決しようとする課題】従来技術では顔を検出する処理の核となる肌色領域抽出の際に、閾値を明示的に与える必要があった。また照明条件の変化によって顔検出に遠した閾値は変化してしまう。上記文献中では画像全体の平均輝度によって閾値を変化させているが、例えば肌色を抽出するための閾値は該当する肌色に依るもので、決して画像全体の平均輝度で定めることができるわけではない。このため、照明条件の変化に対して、より強靭なアルゴリズムが求められている。

【0018】

【課題を解決するための手段】本発明の顔検出装置は、人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗そして明に変化する点と、暗から明に変化する点で挟まれる領域を検出する手段と、該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを取り出して顔候補領域として検出する手段とを有することを特徴とする。

4

【0019】又、本発明の顔検出装置は、人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが暗から明に変化する点と、明から暗そして明に変化する点で挟まれる領域を検出する手段と、該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを取り出して顔候補領域として検出する手段とを有することを特徴とする。

【0020】又、本発明の顔検出装置は、人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗そして明に変化する点と、暗から明に変化する点で挟まれる領域を検出する手段と、該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを顔候補領域として取り出す手段と、標準顔と上記顔候補領域を比較することにより人物の顔を検出する手段とを有することを特徴とする。

【0021】又、本発明の顔検出装置は、人物の顔が撮影されている画像中を、画像の天地方向に調べた時に、明るさが明から暗に変化する点と、明から暗そして明に変化する点で挟まれる領域を検出する手段と、該領域から、該領域の形状特徴と領域内の輝度特徴が人物の目に近い領域を目候補領域として取り出す手段と、上記領域から、上記領域の形状特徴と領域内の輝度特徴が人物の口に近い領域を口候補領域として取り出す手段と、上記目候補領域から2つ、上記口候補領域から1つを重複せずに任意に組み合わせ、人物の顔らしい配置を為している組み合わせを顔候補領域として取り出す手段と、標準顔と上記顔候補領域を比較することにより人物の顔を検出する手段とを有することを特徴とする。

【0022】

【作用】図3（a）に入力画像の例を示す。画像中に人間の正面向きに近い顔が撮影されているとする。

【0023】顔は人それぞれ多様ではあるが、目や鼻、

5

口、顔の形状やその色など、構造上極めて強い制限を持っている。このことから本発明では、特に目と口に注目する。入力画像信号５１を切断面ａ５２や切断面ｂ５３で切断し、その典型的な輝度変化を図４に示す。このように目や口はその周辺に比べて暗くなっていることで簡単な判別をすることができるが、従来技術のように閾値を閾値信号５８のように与えることでは、照明変化に対し頑健なアルゴリズムとは言えない。

【００２４】そこで暗い領域を検出することを止めて、暗い領域近辺即ち図４における輝度が局所的に最も小さくなる輝度極小点５９と、大きくなって行く輝度変曲点ａ６０を検出し、この２点で挟まれる領域を検出することにする。輝度極小点を検出するには画像を一次・二次微分し、その符号がそれぞれ０・＋となる条件により実現できる。輝度変曲点を検出するには同様に画像を一次・二次微分して符号がそれぞれ＋・０となる条件により実現できる。

【００２５】同じように暗い領域検出の代わりに図４における局所的に輝度が小さくなって行く輝度変曲点ｂ６１と、最も小さくなる輝度極小点５９とで挟まれる領域を使用することでも同じ効果を得ることが可能である。但し上記輝度変曲点ａ６０と上記輝度変曲点ｂ６１で挟まれる領域は、双方とも二次微分が０というノイズに弱い条件による検出処理となるために有効ではない。

【００２６】上記領域検出処理は従来技術、例えば特開平５－１８９５６６号公報『エッジ・領域検出方法及び装置』を用いると、閾値を用いずに容易に検出することができる。また周辺に比べて暗い領域を切り取っているだけであることから、白色人種や黄色人種一般に対して有効である。

【００２７】検出された領域は目や口の候補領域となる。この候補領域から、目の候補や口の候補を選び出す。目・口はそれぞれ特徴的な形状やテクスチャをみしていることから、幅や高さとその比、領域内のテクスチャ情報に対して閾やかな閾値を設けることにより、目・口のそれぞれの候補領域を求める。本発明が用いる領域検出処理では輝度の変化するエッジ点を組み合わせて用いているが、一般にエッジ検出処理はノイズに敏感であることから、目・口以外の不必要な自然物体に起因する領域は複雑な形状を成していることから、これらの領域を効果的に削減することができる。

【００２８】更に、人物の顔は通常目二つ・口一つが特異な配列で構成されていることから、既に得ている目候補領域から二つ、口候補領域から一つを重複せずに取り出して、配置を調べて顔らしいか否かのチェックを行う。

【００２９】得られた顔候補の目や口の位置を元にして、原画像から図３（ｂ）のように顔候補画像を切り出し、同図（ｃ）のようなあらかじめ用意した標準顔画像と比較することによって、画像中から顔を切り出すことが可能となる。

6

【００３０】

【実施例】以下に本発明の一実施例を、図１を用いて説明する。

【００３１】図３（ａ）に示した様な、人物の顔を含む画像があり、これを入力画像信号３１とする。本発明である顔検出装置２０は、顔の構成要素を領域として切り出す領域検出装置２１と、顔構成要素領域の大きさや位置関係より顔の候補を出力する顔候補検出装置２２、顔候補を詳しく調べて顔であるか否かを判定する顔判定装置２３の、３つの処理装置から構成される。

【００３２】前記入力画像信号３１を一次微分手段１により、画像上の上から下の方向に一次微分を行い一次微分信号３２を出力する。該一次微分信号を二値化手段２により０で二値化し、一次微分二値化信号３３を出力する。また前記入力画像信号を二次微分手段３により二次微分を行い、二次微分信号３４を出力する。該二次微分信号を二値化手段４により０で二値化し、二次微分二値化信号３５を出力する。

【００３３】論理和手段５により、前記一次微分二値化信号３３と前記二次微分二値化信号３５の論理和をとり目・口第一候補信号３６を出力する。該目・口第一候補信号において連結領域特徴量算出手段６により、連結領域を成している各領域に対して面積、重心位置、領域の縦・横の長さ、また前記入力画像信号をも入力とし、各領域に対して領域内の輝度の平均・分散の領域の特徴量を、領域特徴量信号３７として出力する。

【００３４】該領域特徴量信号を入力とする目第二候補判定手段７が、各領域の面積、領域の縦・横の長さ、輝度平均・分散を調べて、該領域が目らしい領域を判定し、領域の特徴量と合わせて目第二候補信号３８として出力する。同様に前記領域特徴量信号を入力とする口第二候補判定手段８が、各領域の面積、領域の縦・横の長さ、輝度平均・分散を調べて、該領域が口らしい領域を判定し、領域の特徴量と合わせて口第二候補信号３９として出力する。

【００３５】顔候補判定手段９により、前記目第二候補信号から目候補領域を２つ、前記口第二候補信号から口候補領域を１つ、全ての領域が互いに重複しないように選択し、各領域の重心位置を調べて顔らしい配置をしている候補の組を全ての組み合わせについて調べ、顔候補信号４０として出力する。

【００３６】該顔候補信号中の左右の目の候補領域の重心位置を基準に、顔の存在する候補領域を、顔候補領域画像切り出し手段１０によりアフィン変換を用いて切り出し、図３（ｂ）に示した様な、顔候補画像信号４１として出力する。該顔候補画像信号と、図３（ｃ）に示した様な顔標準パターンの間の距離を算出し、ある閾値よりも小さければ、前記入力画像信号の対応する場所に人間の顔が撮影されていると判断し、顔の存在する位置・

7

大きさ・角度を顔信号４２として出力する、顔判定手段１１から構成される。

【００３７】実施例中で述べた上記顔判定装置２３は、最も簡単な例として挙げたものである。該顔判定装置としては従来顔検出装置として用いられてきた様々な技術を適用することが可能であり、上記領域検出装置２１と上記顔候補検出装置２２により顔の存在範囲を効果的に削減することが可能である。また画像中に、顔以外の背景物が余り移っていない状況下であれば、上記領域検出装置２１と上記顔候補検出装置２２の組み合わせだけで、同様の効果を十分発揮することができる。

【００３８】本発明の一実施例を構成する領域検出装置２４を図５を用いながら説明する。前記入力画像信号を一次微分手段１により、画像上の上から下の方向に一次微分を行い一次微分信号３２を出力する。該一次微分信号を二値化手段２により０で二値化し、一次微分二値化信号３３を出力する。論理負手段１２は上記一次微分二値化信号３３を入力とし、論理負を計算して一次微分二値化論理負信号４３を出力する。また前記入力画像信号を二次微分手段３により二次微分を行い、二次微分信号３４を出力する。該二次微分信号を二値化手段４により０で二値化し、二次微分二値化信号３５を出力する。

【００３９】論理和手段５により、前記一次微分二値化論理負信号４３と前記二次微分二値化信号３５の論理和をとり目・口第一候補信号３６を出力する。

【００４０】本領域検出装置２４を図１の領域検出装置２１と置き換えることにより、同様の効果を得ることができる顔検出装置を構成することができる。画像中に、顔以外の背景物が余り移っていない状況下であれば、上記領域検出装置２４と上記顔候補検出装置２２の組み合わせだけで、同様の効果を十分発揮することができる。

【００４１】

【発明の効果】本発明により、該顔の大きさの変化や平行・回転移動、また照明条件の変動に強い、顔検出装置が提供できる。

【図面の簡単な説明】

【図１】本発明である顔検出装置の一実施例の構成図である。

【図２】従来技術である顔検出装置が途中で用いる６つの候補である。（ａ）・（ｂ）・（ｃ）は眼鏡をかけていない場合の候補で、（ａ）が正しく（ｂ）・（ｃ）は目と眉を間違えたものである。（ｄ）・（ｅ）・（ｆ）は眼鏡をかけている場合の候補で、（ｄ）が正しく（ｅ）・（ｆ）は目と眉を間違えた例である。

【図３】（ａ）は入力画像の一例である。（ｂ）は入力画像（ａ）から顔領域をアフィン変換して得た画像である。（ｃ）は典型的な人物の顔画像である。

8

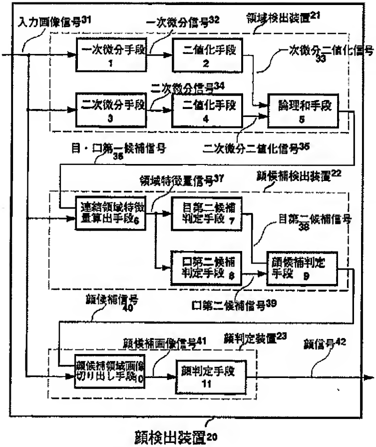
【図４】目・口を横断する輝度信号の典型的な振舞いである。

【図５】本発明である顔検出装置を構成する、領域検出装置の一例である。

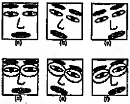
【符号の説明】

１　一次微分手段
２　二値化手段
３　二次微分手段
４　二値化手段
５　論理和手段
６　連結領域特徴量算出手段
７　目第二候補判定手段
８　口第二候補判定手段
９　顔候補判定手段
１０　顔候補領域画像切り出し手段
１１　顔判定手段
１２　論理負手段
２０　顔検出装置
２１　領域検出装置
２２　顔候補検出装置
２３　顔判定装置
２４　領域検出装置
３１　入力画像信号
３２　一次微分信号
３３　一次微分二値化信号
３４　二次微分信号
３５　二次微分二値化信号
３６　目・口第一候補信号
３７　領域特徴量信号
３８　目第二候補信号
３９　口第二候補信号
４０　顔候補信号
４１　顔候補画像信号
４２　顔信号
４３　一次微分二値化論理負信号
５１　入力画像信号
５２　画像切断線ａ
５３　画像切断線ｂ
５４　輝度信号
５５　肌領域ａ
５６　目・口領域
５７　肌領域ｂ
５８　開領域信号
５９　輝度極小点
６０　輝度変曲点ａ
６１　輝度変曲点ｂ

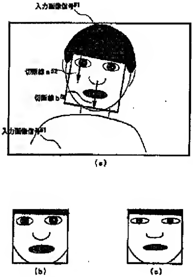
【図1】


入力画像信号31
一次微分信号32
領域検出装置21
一次微分手段 1
二値化手段 2
一次微分二値化信号33
二次微分信号34
二次微分手段 3
二値化手段 4
論理和手段 5
目・口第一候補信号36
二次微分二値化信号35
領域特徴量信号37
顔候補検出装置22
連結領域特徴量算出手段6
目第二候補判定手段 7
目第二候補信号38
口第二候補判定手段 8
顔候補判定手段 9
顔候補信号40
口第二候補信号39
顔候補画像信号41
顔判定装置23
顔信号42
顔候補領域画像切り出し手段10
顔判定手段 11
顔検出装置20

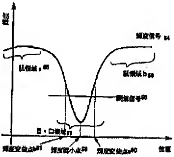
【図2】



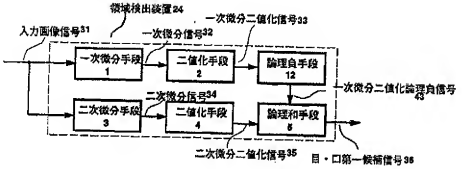
【図3】


入力画像信号51
(a)
(b)
(c)

【図4】



【図5】



フロントページの続き

(51)Int.Cl.[6]	識別記号	序内整理番号	FI		技術表示箇所
G06T 7/00					
G08B 13/196		4234－5G			
		9061－5L	G06F 15/70	460 E	

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]
[Claim 1]A face sensing device comprising:
A point that a luminosity changes from ** to dark and ** when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes from dark to **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means to take out combination which combined one arbitrarily, without overlapping and has succeeded in arrangement appropriate for a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area, and to detect as a face-candidates field.

[Claim 2]A face sensing device comprising:
A point that a luminosity changes from dark to ** when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes to ******** and **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area

and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means to take out combination which combined one arbitrarily, without overlapping and has succeeded in arrangement appropriate for a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area, and to detect as a face-candidates field.

[Claim 3]A face sensing device comprising:
A point that a luminosity changes from ** to dark and ** when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes from dark to **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means which takes out combination which has succeeded in arrangement appropriate for [combine one arbitrarily, without overlapping and] a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area as a face-candidates field, and a means to detect a person's face by comparing a standard face with the above-mentioned face-candidates field.

[Claim 4]A face sensing device comprising:
A point that a luminosity changes from ** tacitly when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes to ******** and **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means which takes out combination which has succeeded in arrangement appropriate for [combine one arbitrarily, without overlapping and] a person's face from two from the

above-mentioned eye candidate area, and the above-mentioned mouth candidate area as a face-candidates field, and a means to detect a person's face by comparing a standard face with the above-mentioned face-candidates field.

[Translation done.]

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Industrial Application]This invention is the art of detecting a person's face out of a picture, and is available to the supervising system of a building, or the visual system of a robot.
[0002]
[Description of the Prior Art]real time parallel cooperative recognition (Osamu Hasegawa.) of the person video for realization of a visual human interface with literature "natural feeling high as an example of a conventional system which detects a person's face from a portrait image Based on Kazuhiko Yokozawa, Mitsuru Ishizuka, and Institute of Electronics, Information and Communication Engineers paper magazine Vol.J77-D-II No.1 Jan (1994)", it explains roughly.
[0003]The algorithm used by this literature extracts the feature shown below as bottom up processing, when video is inputted.
[0004]1. It can set on each point of 2500 points (50x50 points) among a picture among the motion vector 3. picture based on the inter-frame difference of 2500 points (50x50 points) among the average luminance 2. picture of 400 points (20x20 points). the edge feature in the window which the difference between 2500 points (50x50 points) and a setting-out (background) picture set up into the big field 5. picture among the hair based on the ratio of an RGB value, and a beige field 4. picture -- as for the feature extraction processings 3 and 4, setting out of a threshold is needed here. Since the scene which can be adapted if a fixed threshold is given decreases, by the result of the processing 1, the luminosity of the whole picture is distinguished and these thresholds are changed accommodative.
[0005]In the processing 3, the threshold supposing the black hair which a

typical Japanese has, and flesh color is set up. A logical product is taken and the result of the processings 3 and 4 extracts the complexion range which does not exist in a setting-out picture. It uses, when each field is moved, it matches with information and the face which is moving is detected.

[0006]The face model is built on the other hand, using the following conditions as top down processing.

[0007]1. The surface ratio of a hair field and a skin field is less than **30%.

[0008]2. Only the quantity of the range which edge information (an eye, a nose, and a mouth) set up on the complexion range exists.

[0009]3. The above-mentioned edge information has not carried out uniform distribution.

[0010]4. Edge information has gathered in the complexion range upper part comparatively (eyes).

This processing is performed about the field considered that a face exists as a result of bottom up processing, and a detection result is used for the window setting for performing five under bottom up processing.

[0011]As an example of conventional technology, one more. "Target image extraction. for face recognition. using thesub-space classification method (S. based on Akamatsu, T.Sasaki, H.Fukamachi, Y.Suenaga, and MVA'92 pp.465-468(1992)", it explains roughly.)

[0012]When the color picture containing a Japanese person face is inputted, it is supposed by observing this picture in a HSV color space and each YIQ color space that the following knowledge was acquired.

[0013]- On the Q-axis of a YIQ color space, a lip has a high value as compared with the surroundings.

[0014]- The skin field has a characteristic peak on the I-axis of a YIQ color space H axis top of a HSV color space. Then, after carrying out the convert colors of the color picture inputted by RGB to HSV and YIQ, performing a threshold process and detecting a skin field, the above-mentioned feature of a field detects a lip area. Eyes and a mouth make face candidates combination which has accomplished the arrangement appropriate for a face by making the hole in a skin field into an eye candidate.

[0015]What the obtained face candidates were mistaken in eyes and eyebrows, and was detected is contained. The position was normalized by eyes and a mouth, and the face candidates detected by drawing 2 were classified into six, and were shown. (a) is a picture acquired when a detection success is carried out, and (b) and (c) are the examples which

made a mistake that they were eyebrows in one side of eyes, respectively. (d) It is an example of the person to whom –(e) – (f) wore glasses, respectively.

[0016]In this conventional technology, these six normal standard images are prepared, and it outputs noting that it is the face which detected the candidate nearest to (a) or (d).

[0017]

[Problem(s) to be Solved by the Invention]In conventional technology, the threshold needed to be clearly given on the occasion of the complexion range extraction used as the core of the processing which detects a face. The threshold which was suitable for face detection by change of lighting conditions will change. Although the threshold is changed with the average luminance of the whole picture in the above-mentioned literature, it does not depend beige and the threshold for extracting flesh color, for example cannot necessarily be defined [applicable] with the average luminance of the whole picture. For this reason, the tougher algorithm is called for from change of lighting conditions.

[0018]

[Means for Solving the Problem]This invention is characterized by a face sensing device comprising the following.

A point that a luminosity changes from ** to dark and ** when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.

A means to detect a field across which it faces in that it changes from dark to **.

A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.

A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means to take out combination which combined one arbitrarily, without overlapping and has succeeded in arrangement appropriate for a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area, and to detect as a face-candidates field.

[0019]This invention is characterized by a face sensing device comprising the following.

A point that a luminosity changes from dark to ** when inside of a

picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes to ******** and **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means to take out combination which combined one arbitrarily, without overlapping and has succeeded in arrangement appropriate for a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area, and to detect as a face-candidates field.

[0020]This invention is characterized by a face sensing device comprising the following.
A point that a luminosity changes from ** to dark and ** when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.
A means to detect a field across which it faces in that it changes from dark to **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means which takes out combination which has succeeded in arrangement appropriate for [combine one arbitrarily, without overlapping and] a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area as a face-candidates field, and a means to detect a person's face by comparing a standard face with the above-mentioned face-candidates field.

[0021]This invention is characterized by a face sensing device comprising the following.
A point that a luminosity changes from ** tacitly when inside of a picture by which a person's face is photoed is investigated to a top and bottom direction of a picture.

A means to detect a field across which it faces in that it changes to ******** and **.
A means by which a shape facility of this field to this field and the luminosity feature in a field take out a field near a person's eyes as an eye candidate area.
A means by which a shape facility of a described area to a described area and the luminosity feature in a field take out a field near a person's mouth as a mouth candidate area, A means which takes out combination which has succeeded in arrangement appropriate for [combine one arbitrarily, without overlapping and] a person's face from two from the above-mentioned eye candidate area, and the above-mentioned mouth candidate area as a face-candidates field, and a means to detect a person's face by comparing a standard face with the above-mentioned face-candidates field.

[0022]
[Function]The example of an inputted image is shown in drawing 3 (a). It is assumed that the face near for human being's transverse planes is photoed in a picture.
[0023]a face -- a person -- although it is [each] various, the shape of eyes, a nose, a mouth, and a face, its color, etc. have very strong restriction on structure. Especially in this invention, eyes and a mouth are observed from this. The input picture signal 51 is cut in the cutting plane a52 or the cutting plane b53, and the typical luminance change is shown in drawing 4. Thus, although eyes and the mouth can carry out easy distinction because it is dark compared with the circumference, a threshold cannot be said to be a stout algorithm to lighting change by giving like the threshold signal 58 like conventional technology.
[0024]Then, it stops detecting a dark field, the luminosity in dark field neighborhood, i.e., drawing 4, will detect the luminosity minimum point 59 which becomes local the smallest, and the luminosity point of inflection a60 which becomes large and goes, and the field across which it faces by these two points will be detected. Quadratic differential is carried out and primary and the conditions from which the numerals become 0 and +, respectively can realize a picture to detect a luminosity minimum point. A picture is realizable similarly [detecting the luminosity point of inflection] by primary and the conditions from which quadratic differential is carried out and numerals are set to + and 0, respectively.
[0025]Using the field across which it faces instead of similarly dark field detection in the luminosity point of inflection b61 in drawing 4 which

luminosity becomes small locally and goes, and the luminosity minimum point 59 which becomes the smallest can also acquire the same effect. However, since both sides become the detection processing by the conditions that quadratic differential is weak in the noise 0, the field across which it faces in the above-mentioned luminosity point of inflection a60 and the above-mentioned luminosity point of inflection b61 is not effective.

[0026]Described area detection processing can be easily detected, without using a threshold, if conventional technology "edge area detecting method and device", for example, JP,5-189566,A, is used. Since the dark field is only cut off compared with the circumference, it is effective to the white races or general yellow-skinned races.

[0027]The detected field turns into a candidate area of eyes or a mouth. The candidate of eyes and the candidate of a mouth are selected out of this candidate area. Since eyes and a mouth have succeeded in characteristic shape and texture, respectively, it asks for each candidate area of eyes and a mouth by establishing a loose threshold to the texture information in width, height, its ratio, and a field. Although used combining the edge point that luminosity changes, in the field detection processing which this invention uses, Generally, since the field which originates in unnecessary natural objects other than eyes and a mouth since the edge detection process is sensitive to a noise has accomplished complicated shape, it can reduce these fields effectively.

[0028]From comprising arrangement with one usually unique two eyes and a mouth, a person's face is taken out without overlapping one from two from the already obtained eye candidate area, and a mouth candidate area, investigates arrangement, and checks appropriate for a face.

[0029]It becomes possible to start a face out of a picture by carrying out based on face candidates' eyes and the position of a mouth which were acquired, starting a face-candidates picture like drawing 3 (b) from an original image, and comparing with the standard face image prepared beforehand as shown in the figure (c).

[0030]

[Example]One example of this invention is described using drawing 1 below.

[0031]There is a picture containing the face of a person as showed drawing 3 (a), and this is made into the input picture signal 31. The face sensing device 20 which is this invention comprises three processing units, the field sensing device 21 which starts the component of a face as a field, the face-candidates sensing device 22 which outputs the

candidate of a face from face component area size or physical relationship, and the face determining device 23 which questions face candidates in detail and judges whether it is a face.

[0032]The primary differential means 1 performs primary differential in the direction of [lower] for said input picture signal 31 from on on a picture, and the primary differential signal 32 is outputted. Binarization of this primary differential signal is carried out by 0 by the binarization means 2, and the primary differential binarization signal 33 is outputted. The quadratic differential means 3 performs quadratic differential for said input picture signal, and the quadratic differential signal 34 is outputted. Binarization of this quadratic differential signal is carried out by 0 by the binarization means 4, and the quadratic differential binarization signal 35 is outputted.

[0033]By the logical sum means 5, the logical sum of said primary differential binarization signal 33 and said quadratic differential binarization signal 35 is taken, and the first candidate signal 36 of eyes and a mouth is outputted. In this first candidate signal of eyes and a mouth, a connected area to each accomplished field by the connected area feature quantity calculation means 6 Area, A centroid position, the length length and beside a field, and said input picture signal are also considered as an input, and the characteristic quantity of the field of an average and distribution of the luminosity in a field is outputted as the field characteristic quantity signal 37 to each field.

[0034]The second candidate judging means 7 of eyes which considers this field characteristic quantity signal as an input investigates the area of each field, the length length and beside a field, and a luminance average and distribution, and this field judges the field appropriate for eyes, and outputs as the second candidate signal 38 of eyes together with the characteristic quantity of a field. The second candidate judging means 8 of a mouth which considers said field characteristic quantity signal as an input similarly investigates the area of each field, the length length and beside a field, and a luminance average and distribution, and this field judges the field appropriate for a mouth, and outputs as the second candidate signal 39 of a mouth together with the characteristic quantity of a field.

[0035]By the face-candidates judging means 9, an eye candidate area from said second candidate signal of eyes Two. A mouth candidate area is chosen from said second candidate signal of a mouth so that one and no fields may overlap mutually, the group of the candidate who investigates the centroid position of each field and is doing arrangement appropriate

for a face is investigated about all the combination, and it outputs as the face-candidates signal 40.

[0036]On the basis of the centroid position of the candidate area of the eye of the right and left in this face-candidates signal, the candidate area where a face exists is started using affine transformation by the face-candidates region image logging means 10, and it outputs as the face-candidates picture signal 41 as shown in drawing 3 (b). Compute the distance between face standard patterns as indicated to be these face-candidates picture signals to drawing 3 (c), and if smaller than a certain threshold, It comprises the face judging means 11 which judges that human being's face is photoed by the place where said input picture signal corresponds, and outputs the position, size, and angle in which a face exists as the face signal 42.

[0037]The above-mentioned face determining device 23 described in the example is mentioned as easiest example. It is possible to apply various art in which it has been conventionally used as a face sensing device as this face determining device, and it is possible to reduce the existence ranges of a face effectively with the described area sensing device 21 and the above-mentioned face-candidates sensing device 22. If it is under the situation where background things other than a face have seldom moved into the picture, the same effect can be enough demonstrated only in the combination of the described area sensing device 21 and the above-mentioned face-candidates sensing device 22.

[0038]The field sensing device 24 which constitutes one example of this invention is explained using drawing 5. The primary differential means 1 performs primary differential in the direction of [lower] for said input picture signal from on on a picture, and the primary differential signal 32 is outputted. Binarization of this primary differential signal is carried out by 0 by the binarization means 2, and the primary differential binarization signal 33 is outputted. The logic negative means 12 considers the above-mentioned primary differential binarization signal 33 as an input, calculates a logic negative, and outputs the primary differential binarization logic negative signal 43. The quadratic differential means 3 performs quadratic differential for said input picture signal, and the quadratic differential signal 34 is outputted. Binarization of this quadratic differential signal is carried out by 0 by the binarization means 4, and the quadratic differential binarization signal 35 is outputted.

[0039]By the logical sum means 5, the logical sum of said primary differential binarization logic negative signal 43 and said quadratic differential binarization signal 35 is taken, and the first candidate signal

36 of eyes and a mouth is outputted.
[0040]By replacing the special character region sensing device 24 with the field sensing device 21 of drawing 1, the face sensing device which can acquire the same effect can be constituted. If it is under the situation where background things other than a face have seldom moved into the picture, the same effect can be enough demonstrated only in the combination of the described area sensing device 24 and the above-mentioned face-candidates sensing device 22.
[0041]
[Effect of the Invention]By this invention, a face sensing device strong against change of the size of this face, or parallel, a rotation and change of lighting conditions can be provided.

[Translation done.]

* NOTICES *

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1]It is a lineblock diagram of one example of the face sensing device which is this invention.
[Drawing 2]The face sensing devices which are conventional technology are six candidates who use on the way. (a) It is a candidate when -(b) - (c) has not worn glasses, and (a) is [(b) - (c)] correctly mistaken in eyes and eyebrows. (d) It is a candidate when -(e) - (f) has worn glasses, and (d) is the example for which (e) - (f) mistook eyes and eyebrows correctly.
[Drawing 3](a) is an example of an inputted image. (b) is the picture which obtained it from the inputted image (a) by carrying out affine transformation of the face area. (c) is a typical person's face picture.
[Drawing 4]It is typical behavior of the luminance signal which crosses eyes and a mouth.
[Drawing 5]It is an example of the field sensing device which constitutes the face sensing device which is this invention.
[Description of Notations]
1 Primary differential means
2 Binarization means
3 Quadratic differential means
4 Binarization means
5 Logical sum means
6 Connected area feature quantity calculation means
7 The second candidate judging means of eyes
8 The second candidate judging means of a mouth
9 Face-candidates judging means
10 Face-candidates region image logging means
11 Face judging means

12 Logic negative means
20 Face sensing device
21 Field sensing device
22 Face-candidates sensing device
23 Face determining device
24 Field sensing device
31 Input picture signal
32 Primary differential signal
33 Primary differential binarization signal
34 Quadratic differential signal
35 Quadratic differential binarization signal
36 The first candidate signal of eyes and a mouth
37 Field characteristic quantity signal
38 The second candidate signal of eyes
39 The second candidate signal of a mouth
40 Face-candidates signal
41 Face-candidates picture signal
42 Face signal
43 Primary differential binarization logic negative signal
51 Input picture signal
52 Picture cutout line a
53 Picture cutout line b
54 Luminance signal
55 Skin field a
56 Eyes and a mouth field
57 Skin field b
58 Threshold signal
59 Luminosity minimum point
60 Luminosity point of inflection a
61 Luminosity point of inflection b

[Translation done.]

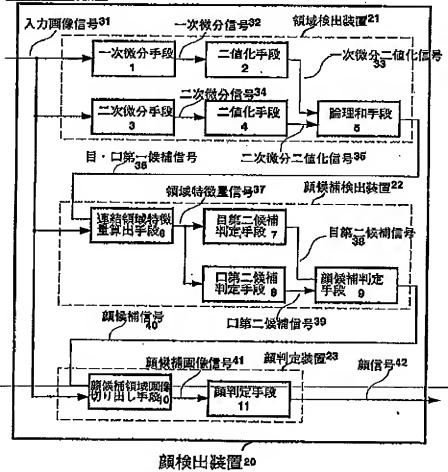
* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.
2.**** shows the word which can not be translated.
3.In the drawings, any words are not translated.

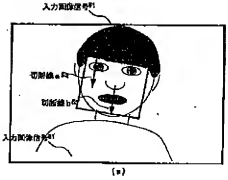
DRAWINGS

[Drawing 1]

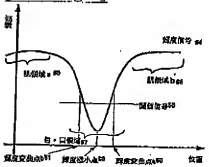


[Drawing 2]

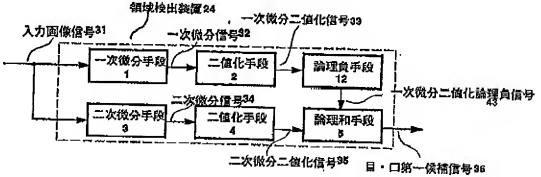
[Drawing 3]



[Drawing 4]



[Drawing 5]



[Translation done.]